Interval estimate of the global energy confinement time during ELMy H-mode in ITER-FEAT, based on the international multi-tokamak ITERH.DB3 dataset

by

Otto J.W.F. Kardaun

Abstract

In the first part of this report, we estimate the technical standard deviation of the thermal energy confinement time in ITER FEAT, based on the public version of the international ITERH.DB3 database. In the second part, it is indicated that the concept of a technical standard deviation is indeed meaningful and had rather not be neglected while concentrating, for example, on point estimation. A log-linear fit for type III ELMy discharges in pure deuterium and an second-order scaling (with quadratic terms on logarithmic scale), are presented. The latter takes, among other effects, roll-over at high density (depending on a plasma shape factor related to triangularity) into account. Both lead to a point estimate of some 3.0 s., corresponding to an H-mode multiplication factor $H_{98y2} \simeq 0.85$ for the present ITER FEAT reference scenario.

In this report we estimate the 'technical standard deviation', see [11] of the global energy confinement-time in ITER FEAT [14] ($R = 6.2$ m, $a = 2.0$ m, $\kappa_a = V/2\pi R\pi a^2 = 1.7$, $\delta_X = 0.5$) for standard ELMy H-mode at the current standard reference scenario, for inductive operation, with plasma parameters $I_p = 15$ MA, $B_t = 5.3$ T, $n_e = 10^{20}m^3$, $P_{L''} = 88$ MW, and $M_{\text{eff}} = 2.5$, $P_{L''}$ denoting an estimate of the net heating power. A number of different approaches are possible to obtain an estimate of this technical standard deviation, some of which have been described in [4]. The present interval analysis is based on the ITERH.DB3v10 (extended) standard dataset for ELMY H-mode, assembled on occasion of the Sorrento IAEA conference, see [7], and electronically available on an internet site from the international database working group maintained at EFDA, see [10]. While T-10 and TUMAN-3 have contributed ELM-free H-mode and TEXTOR radiation-improved L-mode, this database contains at present ELMy H-mode confinement data from 14 tokamaks (ASDEX, ASDEX Upgrade, ALCATOR C-MOD, COMPASS, DIII-D, JET, JFT-2M, JT-60U, PBX-M, PDX, START, TDEV, TCV, TFTR).

As described in [4], the usual statistical error propagation formula leads to too narrow interval estimates for the confinement time in ITER (FEAT). However, it takes relatively well into account the increase of the prediction error as a function of the distance between the ITER reference operating point and the center of gravity of the database. In the first approach considered in this report, we derive an estimate of the multiplication factor c which is to be applied to the classical statistical interval estimate in order to obtain a prediction accounting to some extent realistically for the various sources of variation. Since in ITERH.DB3v10 ELMy H-mode data are collected from 14 tokamaks, instead of from 6 tokamaks in ITERH.DB2, we can now investigate more effectively how well (on average) the confinement time in each machine is predicted when that machine is omitted from the log-linear regression, see Fig. 1. We compare this with the average, classical statistical interval width (using $N_{\mathrm{eff}} = N/4$) in order to estimate the multiplication factor c. The result of analyzing this facet of the problem is presented in Fig. 2. The ITER-like tokamaks AUG, COMPASS, C-MOD, DIII-D, JET, JFT-2M, START are predicted with an average multiplication factor $c = 3.0$. When COMPASS, C-MOD and START are omitted, the average multiplication factor is $c = 2.0$. From this view-point it seems reasonable to take $c = 2.0$ for one and $c = 4.0$ for two standard deviations of a log-linear interval. The case $c = 4$ corresponds to $e^{\pm 0.16}$ for ITER FEAT.

A second aspect is the sensitivity of the point prediction of the thermal confinement time in ITER-FEAT, when all data from one tokamak are shifted by a certain amount, which accounts to some extent for systematic measurement differences between the machines, see [4, Section 6]. The most important effect is produced by JET [8]: shifting $\tau_{E,th}$ from JET by +10% shifts the prediction for ITER FEAT by $+18.5$%, while for all other tokamaks it is less than 3% if we perform OLS regression giving equal weight to all time-slices. Obviously, this large difference between JET and the other tokamaks is partly due to the relatively large fraction of data from JET in the database. Weighting all tokamaks equally, i.e. weighting the time-slices from tokamak j proportional to $1/N_j$, the sensitivity of the ITER FEAT prediction with respect to a +10% shift of $\tau_{E,th}$ from JET is reduced to +9%, for Alcator C-mod and JT-60U the shift is approximately +5%, while shifting DIII-D upwards leads to a downward shift (-5%) for ITER FEAT. For all other tokamaks, the shift is less than 5%. Weighting all tokamaks equally is actually the extreme opposite situation of giving equal weight to each time-slice. A more balanced approach is to consider an intermediate situation, by weighting, for instance, the time-slices from tokamak j proportional to $1/\sqrt{N_j}$. In that case, the sensitivity with respect to JET reduces from $+18.5\%$ to $+14\%$, while the maximum sensitivity with respect to the other tokamaks is still less than 3%, see Fig. 3. From this second view point, we conclude that $+14\%$ appears to be a reasonable, if obviously not a rigorous, estimate of one technical standard deviation of the confinement-time prediction for ITER FEAT.

A third aspect is the empirical distribution of the Jacknife prediction for ITER (FEAT). (In [4] the word Jackknife has been used instead of jacknife, to indicate that an entire tokamak, instead of a time-slice, is left out in turn.) Because of the fairly large number of tokamaks in ITERH.DB3, we can afford now to leave two (or even 3) tokamaks out in turn, without impairing the regression by multi-collinearity. In Fig. 4 the distribution is displayed of the thermal confinement time prediction for ITER-FEAT according to standard log-linear regression when 2 tokamaks are left out in turn, based on the full extended standard ELMy dataset (including ohmic H-mode) of DB3v10, as defined by the criteria described in [7] ($N_{tok} = 14$, $N_{obs} = 2687$). Notice that there are $\binom{14}{2}$ = 91 possible combinations. This figure shows a 'bias' of 9% with respect to the prediction by ITERH-98P(y,2) and, more importantly, an empirical standard deviation of also approximately 6%. One should not be beguilded by the fact that the mean of the Jacknifed estimates is higher than the pont prediction, 3.6 s, from standard ordinary least squares regression, since this mean tends to *overpredict* the true value. The extrapolated (asymptotically bias-corrected) jacknife estimate is somewhat lower than 3.6 s. In [4], arguments were given, that, first, the 'bias' from the Jacknife estimates is not a real correction to the best log-linear point estimate and, secondly, inflating the Jacknife variance with $\sqrt{N_{tok}}$ is not a valid procedure in our situation. The reason for the latter is that deleting an entire tokamak leads to correlations between the Jacknife estimates, which are not negligible. The caveat for applying the standard textbook jacknife method to estimate the variance of the point predictor (and hence the width of the prediction interval), still holds for ITERH.DB3v10, albeit less so, since the number of tokamaks has increased from 6 to 14.

In conclusion, the Jacknife method to estimate the variation in the prediction of the thermal energy confinement time in ITER-FEAT, applied while using log-linear regression and based on the standard choice of (engineering) plasma parameters, yields a value somewhat less than the present technical standard deviation of 14%.

A fourth aspect consists in looking at the variation of the standard log-linear regression for various subsets of the (standard) dataset, while varying the regression weigths as $1/N_j^a$ for $0 \leq a \leq 1$. For $a = 0$ all time-slices are weighted equally and for $a = 1$ all tokamaks are weighted equally. This type of sensitivity analysis was discussed in 1996 just after an ITER Combined Confinement Database and Transport Modelling Workshop in Garching [1], and, applied to a nested breakdown of DB3v5 into 5 different standard subsets, while applying log-linear regression analysis using the traditional set of eight engineering regression variables, resulted in [16, Ch. 2, Appendix A]. An update of this figure, applied to the following subsets of DB3v10, is shown in Fig. 5. (A) The extended standard dataset without ohmic H-mode ($N_{tok} = 13$, $N_{obs} = 2907$) (B) The $D \to D$ subset of (A) ($N_{tok} = 11$, $N_{obs} = 2238$) ; (C) The subset of (A) with type III ELMs or small ELMs ($N_{tok} = 13$, $N_{obs} = 1116$) ; (D) The $D \to D$ subset of (A) with type III ELMs or small ELMs ($N_{tok} = 11$, $N_{obs} = 807$); (E) The DB2 standard ELMy dataset ($N_{tok} = 6$,$N_{obs} = 769$) from the IAEA-1992 conference, see [6, 16].

A fifth aspect is the correction for plasma radiation inside the separatrix. The quantity $P_{L''} = P_{L'} - P_{rad'}$ is the estimated net heat power, and $P_{L'}$ equals $P_L - P_{cx} - P_{ol}$, where $P_L = P_\alpha + P_{oh} + P_{aux}$ is approximately 120 MW for ITER FEAT. The quantity $P_{rad'} = P_{brem} + P_{cycl} + (P_{rad,fb} + P_{rad,bb})/3$ is an estimate of the radiation inside the separatrix, see [2, 15]. For ITERH-98(y,2), $P_{L'}$ is used as a regression variable in the scaling of the thermal energy W_{th}, while for ITER-FEAT $P_{L''}$ is applied. A discussion of this fifth aspect is given in [5].

In the second part of this report, we investigate the influence of log non-linearities in the regression surface. Log-linear models are quite effective in describing the main trend and are rather robust for prediction. Even though log-linear models are quite effective in describing the main trend and are rather robust for prediction, log non-linear models are needed to describe more accurately the regression surface. In [10] is was derived mathematically that the sum of two power-laws, corresponding for instance to a plasma pedestal and a plasma core contribution, somewhat unfortunately, does not have the correct curvature to describe (on a logarithmic scale) the density roll-over, experimentally observed near the Greenwald limit. Special attention has to be paid to the difference in scaling exhibited by the various plasma isotopes (H,D,T). It is tempting to restrict attention, albeit mainly for simplicity, to interaction models, i.e. low-order polynomial models (on logarithmic scale) which can easily be fitted by linear regression programs, which are, for instance, available in SAS and S-PLUS. Even if such models tend to be more suitable for interpolation than for extrapolation, it is interesting to see what they predict for ITER, since they contribute to the interval estimate. One of such scalings, developed on the basis of the extended standard ELMy dataset of DB3v11 (while omitting ohmic H-mode), which is DB3v10 with additional time-slices from AUG (+34), JET (+179) and JT-60U (+33), is

$$\begin{aligned}\ln(W_{th}) \sim\ & 0.9\ln(B_p)+0.1\ln(B_t)+0.37\ln(\overline{n}_e)+0.32\ln(P_{L'}) \\ & +2.95\ln(R)+0.7\ln(\kappa_a)+1.65\ln(a/R)+(2/3)\ln F_{sh} \\ & -0.05\ln(B_p)\times\ln(\overline{n}_e)+0.06\ln(P_{L'}/S)\times\ln(j_p)+\ln(F_{Gr}) \quad (1)\end{aligned}$$

where $B_p = I_p/L$ (with $L = S/(2\pi R)$ the plasma contour length) is the average poloidal magnetic field, B_t the toroidal magnetic field, $\overline{n}_e$ the line-average density, $\kappa_a = V/2\pi R\pi a^2$, $F_{sh} = q_{95}/q_{cyl}$, S the plasma surface area, $j_p = I_p/A$ (with A the plasma cross-sectional area) the current density, and F_{Gr} a density roll-over factor near the Greenwald limit,

$$\begin{aligned}\ln(F_{Gr}) =\ & -0.35\times\ln(\overline{n}_e/n_{Gr})-0.22\times[\ln(\overline{n}_e/n_{Gr})]^2 \\ & +1.5\times[\ln F_{sh}]^2\ln(\overline{n}_e/n_{Gr}) \\ & +1.5\times[\ln F_{sh}]^2[\ln(\overline{n}_e/n_{Gr})]^2 \quad (2)\end{aligned}$$

The units have been chosen such that the logarithms are zero at the ITER FEAT parameters mentioned above. To account for the imbalance in the number of time-slices per device, tokamak j, contributing with N_j time-slices, has been weighted in the regression inversely proportional to the next nearest integer of $2+\sqrt{N_j}/4$. Thermal confinement times from ASDEX were multiplied by 1.15 TAUC92, to account somewhat less stringently than in [6] for the divertor closure, and those from ALC C-mod by 0.85 to account approximately for confinement improvement related to enhanced D-alpha mode compared to type III ELMy H-mode. JFT-2M not being licensed for deuterium injected discharges, and from its similarity to ITER of direct relevance to the estimation of the scaling, the confinement data from this machine have been imputed by a similar type of log-linear interaction scaling as displayed above, based on the mixed isotope dataset ($N \simeq 1100$). The above scaling is based on $N \simeq 860$ time-slices, $D \rightarrow D$, with type-III ELMs or small ELMs. With respect to this dataset, the scaling has an rmse of 14%. For the dataset $D \rightarrow D$ with all types of ELMs with ($N \simeq 2220$) time-slices, the scalings has an rmse of 15%, compared to 16% for the standard simple power law approximation. The interesting point is that the above scaling (as a well as a number of similar log non-linear interaction-type scalings, based either on DB3v10 or DB3v11) yield a point prediction for ITER FEAT of 3.0 s, which is roughly in accordance with the prediction of a simple log-linear scaling based on the pure deuterium ($D \rightarrow D$) subset ($N \simeq 2220$). From Fig. 5 it can be seen that a weighting exponent proportional to N_j^{α} for $1/3 < \alpha < 1/2$, tends to lead to a somewhat increased confinement for ITER FEAT: approximately 3.1-3.2 s.

The main conclusion of this report is that, owing to the extension of the H-mode confinement database from DB2 to DB3v10, see [7], the technical standard deviation of the confinement time prediction for ITER FEAT could be reduced from $\pm 18\%$ to $\pm 14\%$, see Fig. 6.

Secondly, it is remarked that a number of log non-linear interaction models lead to prediction in the lower half of the 95% interval estimate. They are to be treated with prudent caution and provide reason to keep a suitable margin with respect to the point prediction of the energy multiplication factor $Q = P_{fus}/P_{aux}$ in standard, inductive ELMy H-mode. This especially applies to small (type III) ELMs, which are considered to be more suitable than type I ELMs for a prolonged lifetime of the divertor, albeit at the cost of, on average, some 10% to 15% reduction in the energy confinement time. The standard dataset of DB3v5, on which the ITERH-98(y,2) scaling has been based, see [16], actually contains various types of ELMs. For simple power laws, some basic sensitivity analysis related to the distinction between small and large ELMs have been performed during the preparation for [16]. In the meanwhile, some more information on the type of ELM has become available in the database, although not exhaustive and not for all tokamaks and evidence has accumulated that the difference between the two types of ELMs is primarily associated with the log non-linear terms in the confinement scalings, which can be understood from the fact that the type of ELMs is related to edge stability rather than to transport in the confinement zone. type I ELM's occur predominantly in DB3v5, especially so for the JET tokamak. Hence, in conclusion and with some simplification, it seems justified to state that the traditional simple power law scalings based on the mixed isotope confinement ELMy H-mode database, tend to be slanted, within approximately one estimated technical standard deviation, towards non-conservativeness in the confinement time prediction with respect to log non-linear interaction models that take a number of curvatures into account. Especially this holds for divertor-compatible ELMy H-mode (type III ELMs or small ELMs) and calls for some caution when extrapolating from present-day machines to burning plasma experiments.

The sensitivity analysis presented in this report is based upon the ELMy H-mode subset of international ITERH.DB3v10 database, which is at present publicly available. To confirm the size of the interval estimate, an interaction-type log non-linear scaling has been presented based on DB3v10, extended with discharges from AUG, JET and JT-60U, called DB3v11, which led to a point prediction of some 3.0 s. Improved type of H-mode, see e.g. [12], with somewhat peaked density profiles presently not expected in ITER [15], nor pellet enhanced performance discharges [9], nor the operational window associated with type II ELM's [3, 13] have been considered here. In an optimistic vein, these developments contain a promise for enhanced plasma performance at ITER (FEAT), but at present their empirical basis for extrapolation has to be considered as rather scarce.

References

[1] Kardaun, O. and Dorland, W. (1997). Private discussion.

[2] G. Becker. Study of radiation limits in fusion reactor scenarios. *Nuclear Fusion*, 39(7), 1999.

[3] O. Gruber et al. Overview of ASDEX Upgrade results. *Nuclear Fusion*, 41:1369–1390, 2001.

[4] O.J.W.F. Kardaun. Interval estimation of global H-mode energy confinement in ITER. *Plasma Phys. Control. Fusion*, 41:429–469, 1999.

[5] O.J.W.F. Kardaun. On estimating the epistemic probability of realizing $Q = P_{fus}/P_{aux}$ larger than a specified lower bound in ITER. *Nuclear Fusion*, 42(7), 2002.

[6] O.J.W.F. Kardaun for the H-mode Database Working Group. ITER: Analysis of the H-mode confinement and threshold databases. *Proc.*

14th Conf. on Plasma Physics and Contr. Nucl. Fusion Research, Würzburg 1992, 3:251–270, 1993.

[7] O.J.W.F. Kardaun for the International Confinement Database Working Group. Next step tokamak physics: Confinement-oriented global database analysis. In *Fusion Energy 2000 (Proc. 18th Int. Conf. Sorrento, 2000)*, Vienna, 2001. IAEA. CD-ROM file ITERP/04, also available at http://www.iaea.org/programmes/ripc/physics/fec2000/html/node238.htm.

[8] M. Keilhacker, A. Gibson, C. Gormezano, and P.-H. Rebut. The scientific success of JET. *Nuclear Fusion*, 41:1925–1966, 2001.

[9] P.T. Lang et al. Optimization of pellet scenarios for long pulse fuelling to high densities at JET. *Nuclear Fusion*, 42:388–402, 2002.

[10] International H mode Database Working Group. *The International Global H-mode Confinement Database*, 2002 (maintained by K. Thomsen, EFDA). Electronically available at http://efdasql.ipp.mpg.de/HmodePublic.

[11] Kardaun O. Uncertainties in the confinement time prediction for ITER. Technical Report IPP-IR 96/5 1.1, Max-Planck-Institut für Plasmaphysik, 1995. available from http://www.ipp.mpg.de/netreports.

[12] A.C.C. Sips, G.D. Conway, R. Dux, A. Herrmann, J. Hobirk, et al. Progress towards steady-state advanced scenarios in ASDEX Upgrade. *Plasma Phys. Control. Fusion*, 44:A151–A157, 2002.

[13] J. Stober, M. Maraschek, G.D. Conway, O. Gruber, A. Herrmann, A.C.C. Sips, W. Teuterer, H. Zohm, and the ASDEX Upgrade Team. Type II ELMy H-modes on ASDEX Upgrade with good confinement at high density. *Nuclear Fusion*, 41:1123–1134, 2001.

[14] ITER Joint Central Team and ITER Home Teams. Technical Basis for the ITER-FEAT Outline Design and ITER Technical Basis. Technical Report 19 and 24, IAEA, 2000 and 2002.

[15] ITER Joint Central Team and ITER Home Teams. ITER Physics Guidelines. Technical report, Gakunaka, 2001. N-19-FDR-1 (01-06-07) R 0.2.

[16] ITER Expert Groups. ITER Physics Basis. *Nuclear Fusion*, 39:2137–2664, 1999.

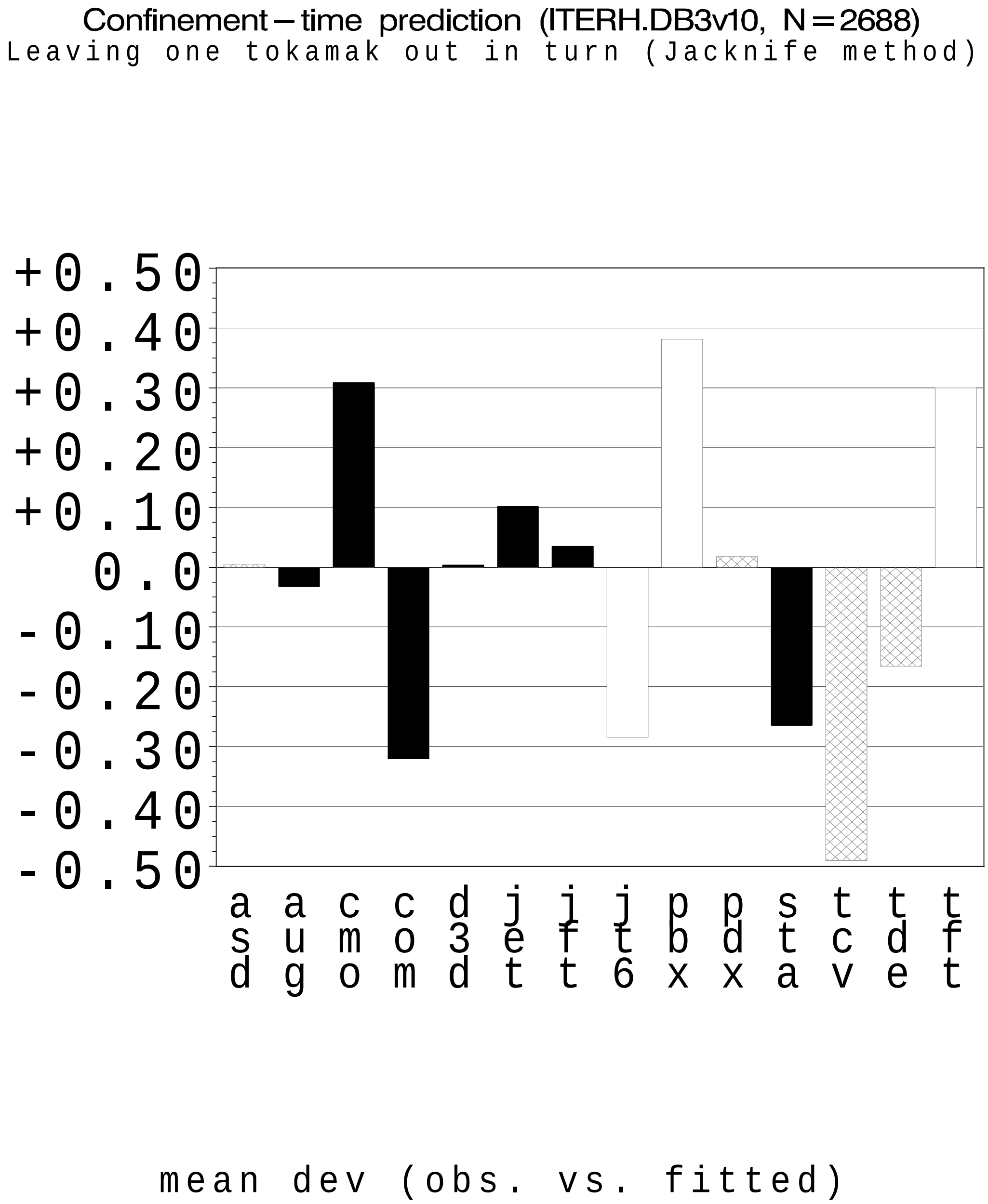


Figure 1: *Average accuracy in predicting the thermal energy confinement time (‘observed minus fitted values’) for standard ELMy H-mode in each of the tokamaks in the extended standard dataset of ITERH.DB3v10 obtained by applying simple log-linear regression to the usual engineering variables. The tokamak for which the confinement time was predicted is not included in the regression data set (‘cross-validation method’).*

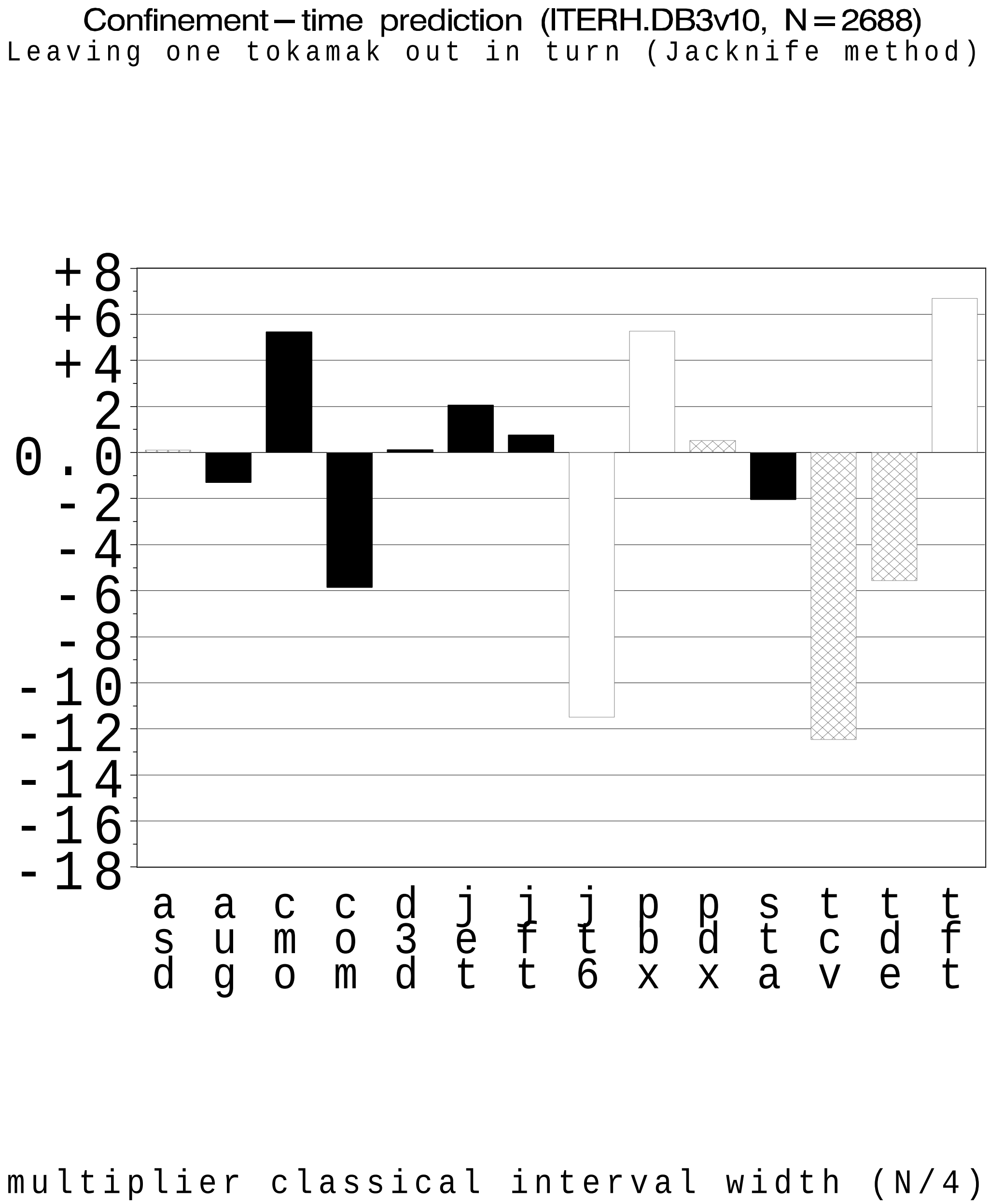


Figure 2: *The plot shows, for each tokamak, the average actual prediction accuracy divided by the average hypothetical, prediction error estimate of the true confinement time for that tokamak, under the assumption that a standard log-linear model holds using $N_{eff} = N/4$ as the 'effective' number of observations. It is noted that the systematic differences between the tokamaks are large with respect to the simple prediction accuracy estimate, however somewhat less so for the tokamaks that have a geometry similar to that of ITER.*

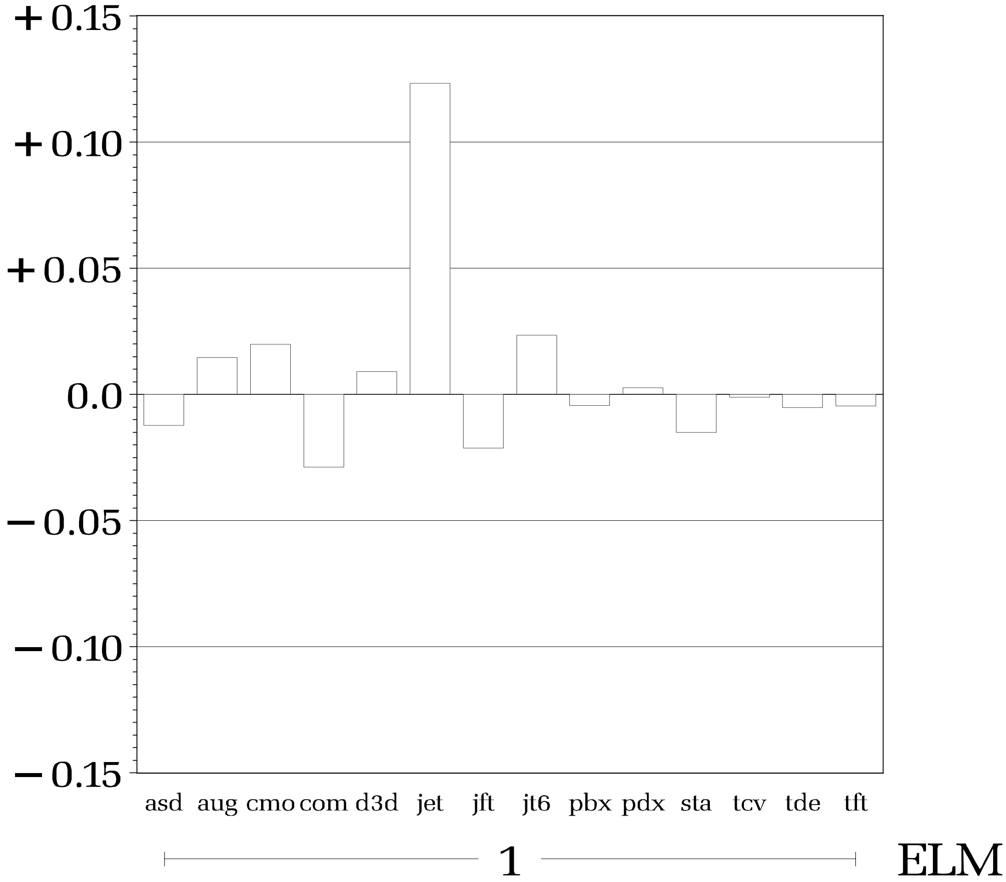
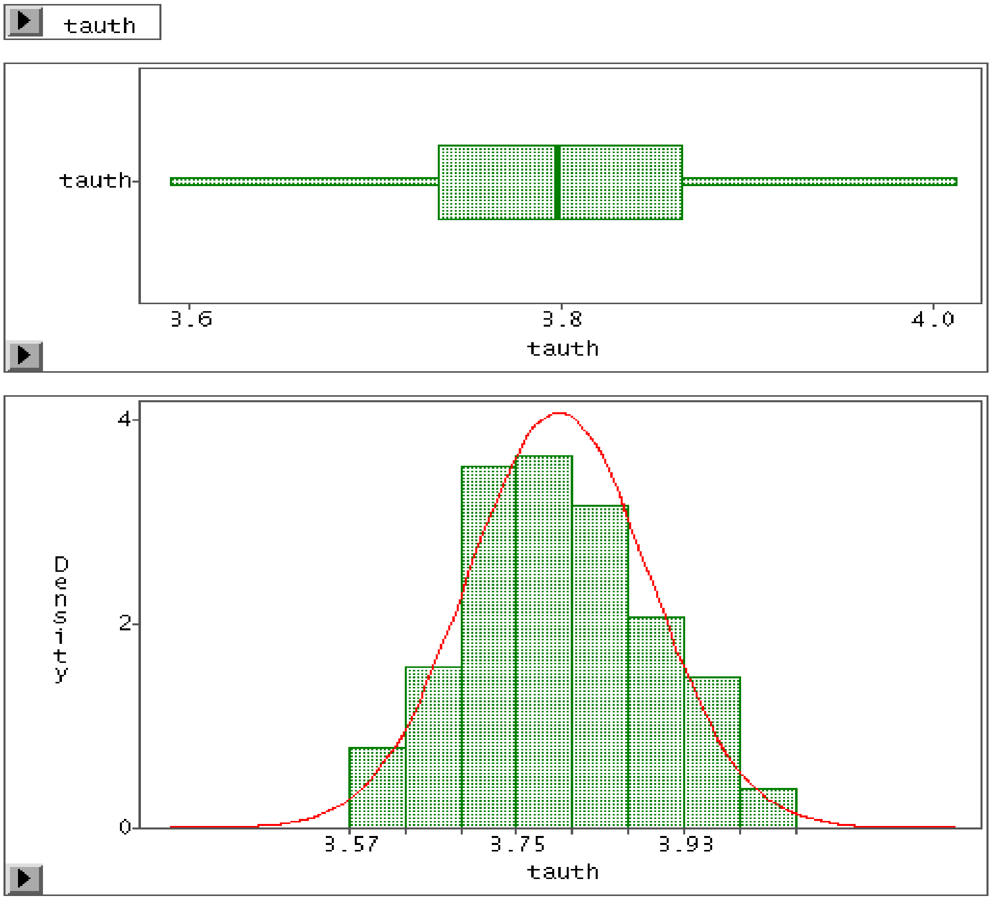


Figure 3: *The sensitivity of the prediction of the thermal energy confinement time in ITER FEAT with respect to shifting the available confinement time data from each tokamak by +10%, while using standard log-linear regression with weights proportional to* $1/\sqrt{N_j}$*, where* N_j *are the number of time-slices contributed by tokamak* j*. The sensitivity of the ITER confinement prediction is highest with respect to the JET tokamak.*



Parametric Density Estimation				
Curve	Distribution	Method	Mean/Theta	:
	Normal	Sample	3.7984	0.097

Moments			
N	169.0000	Sum Wgts	169.0000
Mean	3.7984	Sum	641.9241
Std Dev	0.0978	Variance	0.0096
Skewness	0.1514	Kurtosis	-0.4837
USS	2439.8718	CSS	1.6078
CV	2.5755	Std Mean	0.0075

Figure 4: *Distribution of the predicted confinement times for ITER according to standard log-linear regression, leaving all combinations of pairs tokamaks out in turn. Since the sample average is higher than 3.6 s, the bias-corrected Jackknife estimator leads to a somewhat lower prediction than 3.6 s.*

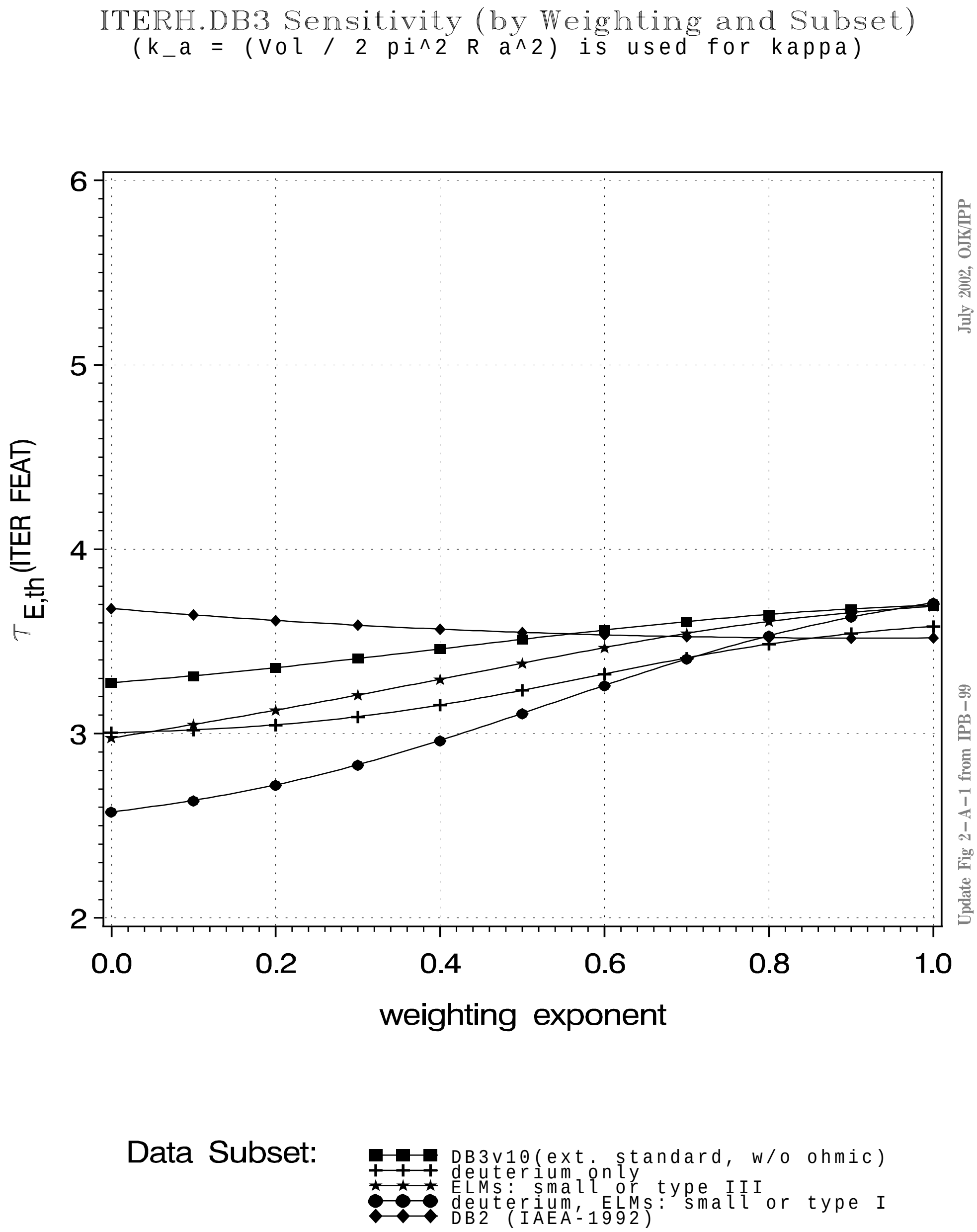


Figure 5: *The plot shows the prediction of the confinement time in ITER FEAT according to weighted regression analysis for five different subsets of ITERH.DB3v10. The regression weights for tokamak j are proportional to N_j^a, for $0 \leq a \leq 1$, where N_j is the number of time-slices contributed by tokamak j.*

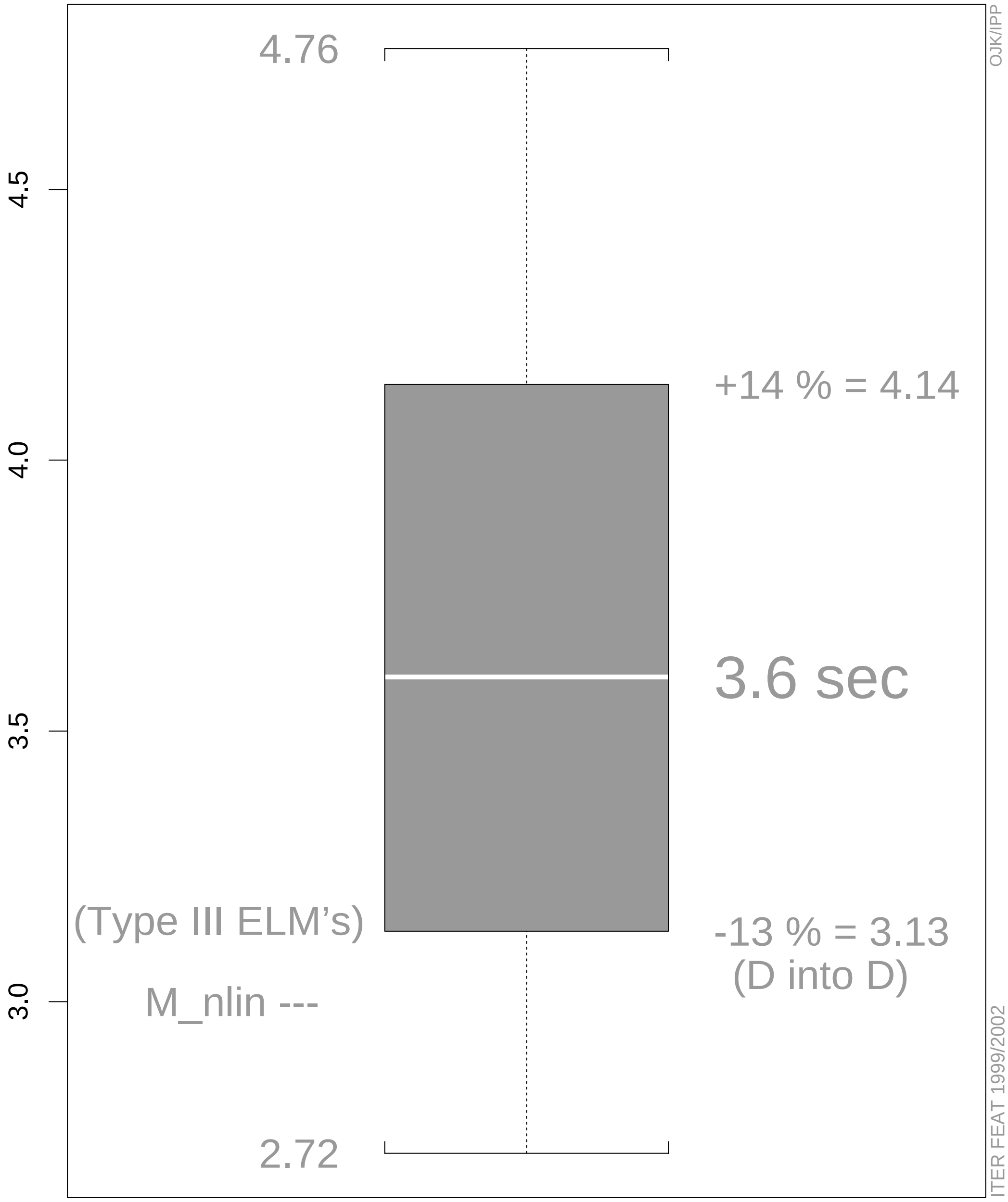


Figure 6: *Interval estimate of the confinement time in ITER FEAT, based on sensitivity analysis of the ITERHDB3v5 and ITERH.DB3v10 databases with (predominantly) type I ELMs.*